DOCUMENT RESUME

ED 381 187 JC 950 165

AUTHOR Haden, Dan; Wells, Sue
TITLE Evaluation Bias in Prior Learning Assessment Challenge Processes.
INSTITUTION Centennial Coll., Scarborough (Ontario). Centre for Instructional Development.
SPONS AGENCY Ontario Council of Regents, Toronto.
PUB DATE [94]
NOTE 16p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Age Differences; *Cultural Differences; *Evaluation Methods; Evaluation Research; Foreign Countries; *Prior Learning; *Student Characteristics; Technical Institutes; Test Bias; Test Format; Two Year Colleges
IDENTIFIERS *Centennial College of Applied Arts and Techn ON

ABSTRACT

In 1993, the Centre for Instructional Development at Centennial College, Ontario, Canada received funding to investigate the effects of student diversity on prior learning assessment evaluation processes and to develop guidelines to respond to any effects demonstrated. A review of the literature identified six barriers to equitable evaluation: alienation; diminished self-confidence; slow reaction time; impaired vision and hearing; English language difficulty; and learning style. In addition, the literature suggested eight strategies to overcome the identified barriers: neutral language in evaluation instructions; scrutiny to assure bias-free test content; definition of clear outcomes; flexible evaluation to accommodate learning and culture and the use of various means of score analysis; self-administered tests and the provision of evaluation options to students; regular formative evaluation; extending peer-tutoring and other group models to the evaluation process; and performance-based assessment evaluating skills in a natural setting. A survey was conducted at Centennial College to identify evaluation methods that students and faculty felt provided the most fair and equitable opportunity to illustrate students' knowledge and skills. The survey obtained responses from 1,542 continuing education students, 1,688 full-time students, and 156 faculty. The study revealed that some cultural groups had a wide range of experiences with evaluation formats while others had only limited exposure; female students and students under 25 were more experienced with all formats than male students and older students; students who spoke only English at home had more experience with all formats; and students completing their highest educational level 1 to 3 years prior to the survey experienced a broader range of testing than students completing their highest educational level more than 3 years prior to the survey. Based on the literature review and survey findings, faculty were urged to define learning outcomes; carefully consider the evaluation format; examine contents closely; link things to connect personal experience with the challenge process; and provide formative features. (Contains 17 references.) (KP)

ED 381 187

Evaluation Bias in Prior Learning Assessment

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Dan Haden
Sue Wells

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S. Wells

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

Centre for Instructional Development
Centennial College

BEST COPY AVAILABLE

JC 950 165

EVALUATION BIAS IN PRIOR LEARNING ASSESSMENT CHALLENGE PROCESSES

by Dan Haden & Sue Wells

Introduction

In 1993, The Centre for Instructional Development (CID) at Centennial College received funding from the Ontario Council of Regents to investigate the effects of student diversity on prior learning assessment evaluation processes and to develop responsive guidelines for any effect(s) demonstrated (the 'project'). The literature review and demographic analysis conducted as part of the project resulted in a 'Responsive PLA Challenge Processes' package consisting of; an interactive computer presentation entitled 'The Prior Learning Assessment Challenge Process Adventure Tour'; a corresponding 'hard copy' booklet; a consolidation of 'Just the Facts'; and a one page 'Summary Sheet'. The object of the different presentations was to appeal to the various learning styles of users. This article discusses the project, presents some of its conclusions and limitations, and makes suggestions for future research.

Diversity

While the fact of student diversity within the college system is being increasingly acknowledged, it appears from issues which are arising that we are only beginning to understand how such diversity is reflected in our educational responsibilities and routines. According to Grosz (1990 p.17) our lack of understanding may originate in how our particular race and socioeconomic status affects our perceptions, such that the perceptions of the majority of educators "... are so different from those of minority students they sometimes struggle to help that they simply cannot perceive the nature of the problem." The ensuing biases are so pervasive and influencing that they have also become rooted in our "... historical, political, social, economic and legal structure[s] ..." (Central Region Project, 1993 p.30). Within teaching responsibilities, these biases clearly affect evaluation systems and, consequently, the students who participate in them (Smith, 1989 p.9). When, in turn, educators witness evaluation differences in students, the tendancy is to "... assume the institution's perfection and the students' incompetence." (Smith, 1989 p.65) and move towards 'special needs' remediation (Wells, 1994 'Literature Review' p.2).

Literature Review - Barriers to Equitable Evaluation

One of the first tasks of the project was to conduct a literature review to determine "... what types of barriers may exist in the evaluation tool or its administration which may prevent access to equal opportunity for the learner to fairly demonstrate his or her learning." (Ontario Council of Regents, 1993 p.13). Within the community college system, equal opportunity to success is an issue of education equity (Vision 2000 p.30).

Generally, six factors were identified as significant characteristic barriers. These included alienation, self-confidence, reaction time, vision and hearing, language, and learning style. They were identified in the following ways:

1. References to alienation and isolation occur frequently in the related literature,

especially literature respecting minority students, women, disabled students and adult learners (Smith, 1989 p.iii). Immigrant students seem particularly vulnerable where their feelings of alienation are experienced within demands imposed by new environments (Kiang, 1992 p.101).

2. Diminished self confidence was also a barrier, especially for students with english as a second language or whose educational background was other than based on the eurocentric model.

3. Adult students and those with certain physical or language distinctions may have slower reaction times which adversely affect learning processes (Cross, 1988 p.132).

4. Any extent of visual or hearing impairment may affect the learning and evaluation process (Cross, 1988 p.135).

5. English language difficulty can challenge learners during evaluation processes.

6. Learning style is one of the most influential barriers, especially for minority and learning disabled students. Knowing that students learn differently has led some to conclude that students should also be evaluated differently (Smith, 1989 p.62; Rodriguez, 1991 p.7).

Literature Review - Strategies to Reduce Barriers

From the literature review the project compiled the following eight types of strategies to overcome the identified barriers:

1. Evaluation instruments should be comprised of neutral language, avoiding the use of slang, idioms and colloquialisms. This does not negate reference to culture (Byers, 1993 p.72), but may mean accomodating the student through translation (Byers, 1993, p.82; Santiago, 1992 p.31), oral examination (Byers, 1993 p.83), or evaluating content separately from language (Santos, 1986 p.17).
2. Experience from American truth-in-testing legislation and especially the Illinois golden rule bias reduction principle substantiates that certain evaluation items may be culturally biased or contain culturally biased distractors. (Weiss, 1987 p.5). This emphasizes the need for scrutiny to assure bias free content.

3. "Defining clear outcomes [was] an important strategy discussed frequently in the literature ..." " Evaluation criteria should be explicit ... [and] obvious" (Wells, 1994 'Literature Review' p.5). Performance should be based on valid indicators of learning (Smith, 1989 p.62) which are also suficiently rigorous (Rodriguez, 1991 p.8).

4. Flexibile evaluation can accomodate learning and culture (Smith, 1989 p.65; Rodriguez, 1991 p.8), and measure quality in light of diversity (Smith, 1989 p.65). As well as reducing barriers to fair evaluation, this could increase validity through score analysis across various evaluation methods (Byers, 1993 p.82).

5. Increasing the locus of control of the learner over the evaluation will increase the learner's self-confidence. This might be achieved through self-administered tests and

permitting evaluation options (Byers, 1993 p.82).

6. Regular formative evaluation supports positive self-esteem and otherwise reduces barriers to fair evaluation by promoting goal directed behaviour (Guba & Lincoln, 1985 p.31). This is especially effective for students whose self-esteem is negatively affected by the uncertainty of goal deferrment (Blum, 1992 p.25).

7. Consideration should be given to creatively extending supportive peer-tutoring and other group models to the evaluation process (Rodriguez, 1991 p.12) to reduce student feelings of alienation. In applying such models however, instructors must be careful not to promote learned helplesness (Blum, 1992 p.25).

8. "Performance based assessment ... [which] evaluates skills in a natural setting" (Wells, 1994 'Literature Review' p.6) supports both the adult student's expectations that study will validate experience (Smith, 1989 p.67) and the particular strength of high risk students to work with tangible problems (Blum, 1992 p.37).

Survey on Evaluation Methods

Concluding, in part, that the literature "did not speak to the unique issues that various diversities may bring to both the evaluation process ... and ... specific types of evaluation[s]" (Wells, 1994 'Literature Review' p.7), the project recommended that a survey " ... to identify evaluation methods participants felt provided them with the most fair and equitable opportunity to illustrate knowledge and skills" (Wells, 1994 'Literature Review' p.7) be designed and distributed to the current Centennial student body. Anticipating the possibility that allowing students to indicate their most preferred evaluation method might be interpreted by some as validation of a student propensity to avoid a more difficult evaluation choice, the project premised that the reason for any particular student's choice may, in fact, lie in such factors as the student's age, gender, experiences, etc., or otherwise their evaluation history. If students feel they are better able to exhibit knowledge and skills in a particular evaluation method, it might be due to something in their experiences which allows them to more successfully manage that method.

In consultation with the Ontario Anti-Racism Secretariat and pursuant to relevant legislation, including the Freedom of Information and Protection of Privacy Act R.S.O. 1990, c.F.31, a survey was designed which related the evaluation experiences and preferences of the then current student and faculty population.

The survey was administered between March 19th and 21st, 1994. 1,542 continuing education students, 1,688 full time students and 156 faculty responded. Although results were compiled for the college as a whole, many results were also tabulated separately for the day and continuing education schools.

The survey analysed 7 different major demographic groupings, namely; age, gender, highest level of education, country of education, length of time since education completed, language, and cultural background. Each grouping was asked to respond to questions about their experiences with a list of evaluation formats including essay, short answer, fill in the blank/completion, true or false, multiple choice, demonstration or performance assessment, product assessment and, interview or oral examination. Specifically, students were asked to

identify formats; they had experienced; they thought difficult or unfair and; for which they thought a lack of skill in english or a special need would cause a lower mark. The survey did not ask respondents to identify the reasons for their preferences.

In addition, faculty were asked to identify evaluation formats they; had used; thought were not suited to their course; thought required too much time and effort to use, and; thought would especially challenge students with lack of skill in english.

Since the survey was comprehensive, only those portions relating to the conclusions of the project are reported here.

Survey Results

When the survey data was independantly tabulated results interestingly identified relationships between demographic groupings and experience with certain evaluation format(s). According to the 'Just the Facts' portion of the project, students' reports indicated that;

(a) "Some cultural groups ... [had] a wide range of experiences with evaluation formats, while others ... [had only] limited exposure (Wells, 1994 p.13);

(b) female students and students under 25 were more experienced with all formats than were male students and those over 25;

(c) almost hierarchically, students whose language at home was english only had more experience with all formats than those who spoke english and another language at home. In turn, students who spoke english and another language at home had more experience with all formats than those who did not speak english at home, and;

(d) students completing their highest education level 1-3 years ago or within Canada experienced "a broader range of testing" (Wells, 1994 p.13) than students completing their highest education level more than 3 years ago or outside Canada.

Given these relationships, the project 'Summary Sheet' concluded that processes used by faculty in selecting evaluation format(s) may be influenced by unique past experiences and may wrongly assume "... that the student already possesses skills and abilities consistent with that format. These may have little or nothing to do with the content that is being assessed." (Wells, 1994) Such individual assumptions may bias evaluation results.

Strategies

Based on the literature review and survey, the project recommended 5 strategies faculty could use in selecting evaluation formats which would take the research results into consideration.

The first of these strategies, 'Defining Learning Outcomes' takes advantage of recent trends in the community college system. Learning outcomes "describe performances that demonstrate that significant learning has been verified and achieved ... (Wells, 1994 'Just the Facts' p.14). When determined before the course content is delivered, learning outcomes can

be weighted by assigning a percentage mark value to the respective worth of the outcome, and designed to measure learning at any of the traditional taxonomical levels. This clarifies what is being evaluated and guides the design of the evaluation instrument.

The next strategy, 'Carefully Considering the Format' asks the reader to consider several options, including; allowing the student to select from several formats; using more than one format within a test; selecting a format the student has previously experienced; and using a format the student feels will best demonstrate their skill/knowledge.

'Examining the Content Closely' provides a list of questions which help determine the extent to which an evaluation format may bias. The questions are designed to alert the reader to consider; offensive and/or unfamiliar references; stereotypical representations; such language peculiarities as slang, idioms, reduced forms and simple sentences; multiple concepts, and; illustrations.

'Linking Things Together' helps students "connect personal experiences with the challenge process" (Wells, 1994 'Just the Facts' p.18). According to the project this is of special importance for students with the characteristics of diversity. Linking things together can be achieved by giving students opportunity to practice the type of evaluation(s) they will experience, providing tips for studying - perhaps by referring the student to available resources for improving studying, or even by connecting the student with other students who have completed the evaluation process.

The last suggestion offered by the project is 'Providing Formative Feedback'. This involves communicating with the student about their progress, usually through continuing evaluation which may or may not form part of the student's grade. This reduces feelings of alienation and isolation.

Conclusion

In conclusion, while the project accepted uniqueness of personality and experience as a certain factor in the delivery of curriculum content, the precise effect of various diversities on the range of available evaluation processes remains to be better understood. This study attempts to heighten awareness and contribute to a greater understanding of the potential for evaluation bias - as may be centered in the learner, the teacher and, most importantly, in the diverse differences between them, and to provide a foundation for a more scientific understanding from which future research can take direction. This study is only an initial step, and is subject to such limitations as are above-mentioned (including that respondents were not asked to explain the reason(s) for their responses), and as may be scientifically concluded from further research.

Specifically, future research could;

(a) more precisely clarify and define the effect of diversity(ies) on individual evaluation formats;

(b) determine whether the fact of evaluation bias is generalizable to other evaluation scenarios, e.g. classroom, employee and course evaluation;

(c) determine the extent to which the current understanding of evaluation bias in other disciplines such as human resources management are incorporable or generalizable to knowledge/skill evaluation;

(d) propose other and/or more effective strategies for reducing the possibility or effect of evaluation bias;

(e) contribute to a greater understanding and propensity toward fair and equitable evaluations generally;

to name a few.

REFERENCES

Blum, M.E. and others.(1992) Developing Educational Programs for the High Risk Secondary School and College Student. Columbia University.

Byers, P. (1993) Evaluation in a Culturally Diverse Classroom. In 'Pro Diversity: A Guide for Achieving Equity in the Classroom'. George Brown College: Toronto, ON.

Central Region Project (1993) Pro Diversity: A Guide for Achieving Equity in the Classroom. George Brown College: Toronto, ON.

Cross, P. (1988) Adults as Learners. Josey-Bass: San Francisco, CA.

Grosz, K. (1990) The Challenge of Cultural Diversity in the Classroom. The Forum, Academic Senate for Community Colleges. v.7, Summer, 1990. California: Sacramento.

Guba, E. and Lincoln. Y. (1985) Effective Evaluation, Improving the Usefulness of Evaluation Results Through Responsive and Naturalistic Approaches. Josey Bass: London.

Kiang, P.N.(1992) "Issues of Curriculum and Community for First-Generation Asian Americans in College." New Directions for Community Colleges. No. 80, Winter, 1992.

Ontario Council of Regents (1993) PLA and Ontario's Culturally Diverse Population - Prior Learning Training. A workshop sponsored by the Ontario Council of Regents, May 10-14, 1993. Toronto, ON.

Rodriguez, A. (1991) Multi-Cultural Education: Some Considerations for a University Setting.

Santiago, R.L. (1992) The Education of Hispanic Americans. The Issues and Language and Culture: Proceedings of a Symposium Convened by the Center for Applied Linguistics. October 5, 1992. Washington, D.C.

Santos, (1986) Professor's Reactions to the Academic Writing of Non-Native-Speaking Students. California State University: Los Angeles.

Smith, D.G. (1989) The Challenge of Diversity: Involvement or Alienation in the Academy? ASHE-ERIC Higher Education Report No. 5, 1989.

Ministry of Colleges and Universities (1990) Vision 2000: Quality and Opportunity. Toronto.

Weiss, J. (1987) Truth-in-Testing & thr Golden Rule Principle: Two Practical Reforms. Paper presented at the Annual Meeting of the National Council on Measurement in Education: Washington, D.C. April 21-23, 1987.

Wells, S.B. (1994) Review of the Literature: Prior Learning Assessment Challenge Process Project. Centennial College: Scarborough, ON.

Wells, S.B. (1994) Just the Facts: Prior Learning Assessment Challenge Process Project. Centennial College: Scarborough, ON.

Wells, S.B. (1994) Summary Sheet: Prior Learning Assessment Challenge Process Project. Centennial College: Scarborough, ON.

Dan Haden is an instructor in legal skills in the School of Continuing Education at Centennial College. He is author of the Ontario Title Searching Handbook: Carswell, and the Ontario Family Law Procedure Handbook: Carswell.

Susanne Wells is chairperson for the Child Studies Department at Centennial College. She is co-author of 'Learner Support Strategies: An Integrated Approach', Centennial College (ERIC).

APPENDIX 'A'

SUMMARY OF SURVEY

EXECUTIVE SUMMARY

- The Centennial 'Survey Report on Evaluation Methods' was administered during March 19th-21st, 1994, to 3230 students and 156 faculty from 188 classes.

- Demographic characteristics that illustrate the diversity of the students surveyed are as follows: 32% were *'30 years of age or older'*; 32% had *'already graduated from a College or University program'*; 25% had received their *'highest level of education outside Canada'*; 14% *'do not speak English at home'*; 30% *'usually speak English at home, but also spoke another language'*; the largest non-English language group was *'Chinese'* (23%); the largest non-British/Canadian culture group was *'Caribbean origin'* (17%).

- Notable differences in the demographics reported by *'Day'* and *'CE'* students surveyed, were as follows: Age *'between 20-24'* (*'Day'* - 57%, *'CE'* - 22%); *'Highest level of education - College or University graduate'* (*'Day'* - 16%, *'CE'* - 49%); Obtained *'highest level of education outside Canada'* (*'Day'* - 15%, *'CE'* - 25%); Obtained *'highest level of education more than 3 years ago'* (*'Day'* - 32%, *'CE'* - 67%).

- BEFORE COMING TO CENTENNIAL the following groups reported more experience with all types of testing than their immediate counterpart (i.e. *'female'* - *'male'*): those *'<25 years of age'*; *'female'*; *'grade 13 or more'*; *'educated in Canada'*; *'highest level of education 1-3 years ago'*; *'speak English always at home'*.

- WHILE AT CENTENNIAL the following groups reported more experience with all types of testing than their immediate counterparts: those *'<25 years of age'*; *'female'*; *'educated in Canada'*; *'highest education 1-3 years ago'*; *'speak English always and only'*.

- WHILE AT CENTENNIAL *'Day'* students reported more experience with all types of testing than their *'CE'* counterparts. The largest difference was seen in **essays** where 76% of *'Day'* students said they had been tested by an **essay** compared to 41% of *'CE'* students. (Smallest - **product assessment**: 29% *'Day'*, 21% *'CE'*).

- Generally, students expressed a dislike for the **essay** method of evaluation and a high percentage stated that a *'lack of skills in English'* would likely result in lower marks when the essay method was used. Faculty in *'CE'* tended to agree with this while *'Day'* faculty were more divided in their opinion.

- Students and faculty gave significantly different responses to how much the **demonstration** and **product assessment** methods of evaluation were being used.

- **Short answer, true or false** and **multiple choice** methods of evaluation were well supported by students and faculty.

- Generally, students expressed a dislike for the **interview** method of evaluation and 38% felt it was *'unfair'*. Similarly 59% of the students felt this method of evaluation would work against those with *'special needs'*.

- Although there were significant differences in how demographic groups responded to evaluation methods, the most remarkable result was that students were almost always unanimous in the way they ranked evaluation methods. For instance, when asked which of the evaluation methods would likely give lower marks if *'skills in English were lacking'*, all demographic groups ranked **essay** first, **interview** second, and all other methods a distant third.

INTRODUCTION

- The Centennial 'Survey Report on Evaluation Methods' was administered during March 19th-21st, 1994, to 3230 students and 156 faculty from 188 classes.
- The survey was administered to 1542 *'CE'* students out of a potential of 2884 for a return rate of 53%. The survey was administered to 1688 *'Day'* students out of a potential of 4003 for a return rate of 42%. Overall, 3230 students completed the survey out of a potential 6887 for a return rate of 47%.
- The survey was administered to 104 *'CE'* classes out of a potential 146 for a return rate of 71%. It was also administered to 84 *'Day'* classes out of a potential 166 for a return rate of 51%. Overall, 188 classes completed the survey out of a potential 312 for a return rate of 60%.
- The survey was administered to 83 *'CE'* faculty out of a potential 146 for a return rate of 57%. It was completed by 73 out of 166 *'Day'* faculty for a return rate of 44%. Overall, 156 out of 312 faculty completed the survey for a return rate of 50%.
- It should be noted that a number of faculty did not complete a survey even when students for that class did.
- For purposes of analysis 42 demographic groupings were composed as shown on page v. All student responses to question 9 were checked against these groupings to see if there were any relationships. The size of groups 19, 21 and 25 was considered too small for meaningful comparisons.
- When making comparisons, two categories were often presented in brackets to demonstrate the range of variation. All other categories appeared somewhere between these numbers.
- During the analysis it became apparent that the *'CE'* and *'Day'* breakdown would be useful. In particular, it helped in the analysis of essay type questions.